

PRACTICING DATA SCIENCE WITH MS EXCEL AND PYTHON				
Course Code:		AD2601-1	Course Type:	PCC Lab
Teaching Hours/Week (L: T: P):		0:0:2	Credits:	01
Total Teaching Hours:		0+0+26	CIE + SEE Marks:	50+50
Prerequisite:		CS1002-1, CS1005-1		
Teaching Department: Artificial Intelligence & Data Science				
Course Objectives:				
1.	To understand the advanced features of MS Excel			
2.	To apply the features of MS Excel for data analysis and related applications.			
3.	To learn the libraries of python that are used for data analytics.			
4.	To develop programs for data analysis using python libraries			
List of Experiments				
Part- A: MS- Excel				
1	Exploring data using charts and graphs& Calculating summary statistics: <ul style="list-style-type: none"><li>Create a histogram of a dataset to visualize its distribution.</li><li>Create a scatter plot to see the relationship between two variables.</li><li>Create a line chart to track changes over time.</li><li>Calculate the mean, median, and mode of a dataset.</li><li>Calculate the standard deviation and variance of a dataset.</li><li>Calculate the quartiles and interquartile range of a dataset.</li></ul>			
2	Building predictive models: <ul style="list-style-type: none"><li>Use the LINEST function to perform linear regression analysis on a dataset.</li><li>Use the FORECAST function to make predictions based on a linear regression model.</li><li>Use the TREND function to create a trendline for a dataset.</li></ul>			
3	Creating pivot tables: <ul style="list-style-type: none"><li>Create a pivot table to summarize a dataset by different categories.</li><li>Create a pivot chart to visualize the data in a pivot table.</li><li>Use slicers to filter data in a pivot table.</li></ul>			
4	Data cleaning and preprocessing: <ul style="list-style-type: none"><li>Use the TRIM function to remove leading and trailing spaces from text data.</li><li>Use the CLEAN function to remove non-printable characters from text data.</li><li>Use the SUBSTITUTE function to replace specific text in a dataset.</li></ul>			
5	Hypothesis testing: <ul style="list-style-type: none"><li>Use the T.TEST function to perform a t-test on two samples.</li><li>Use the ANOVA function to perform an analysis of variance on multiple samples.</li><li>Use the CHISQ.TEST function to perform a chi-square test on categorical data.</li></ul>			
Part- B: Python				
6	Numpy Library: <ul style="list-style-type: none"><li>Create a numpy array from a list, a tuple with float type.</li></ul>			

	<ul style="list-style-type: none"> <li>• Python program to demonstrate slicing, integer, and Boolean array indexing.</li> <li>• Write a python program to find min, max, sum, the cumulative sum of an array.</li> <li>• Write a Python program to demonstrate use of ndim, shape, size, dtype.</li> </ul>
7	Numpy Library: Linear Algebra <ul style="list-style-type: none"> <li>• Write a python program to find rank, determinant, and trace of an array.</li> <li>• Write a python program to find eigenvalues of matrices.</li> <li>• Write a python program to find matrix and vector products (dot, inner, outer, product), matrix exponentiation.</li> <li>• Write a python program to solve a linear matrix equation, or system of linear scalar equations.</li> </ul>
8	Pandas Library <ul style="list-style-type: none"> <li>• Write a python program to implement Pandas Series with labels.</li> <li>• Create a Pandas Series from a dictionary.</li> <li>• Creating a Pandas DataFrame.</li> <li>• Write a program which make use of following Panda's methods.               <ol style="list-style-type: none"> <li>a. describe ()</li> <li>b. head ()</li> <li>c. tail ()</li> </ol> </li> </ul>
9	Pandas Library: Selection
10	<ul style="list-style-type: none"> <li>• Write a program that converts Pandas DataFrame and Series into NumPy. Array.</li> <li>• Write a program that demonstrates the column selection, column addition, and column deletion.</li> <li>• Write a program that demonstrates the row selection, row addition, and row deletion.</li> <li>• Get n-largest and n-smallest values from a particular column in Pandas dataframe.</li> </ul>
11	Pandas Library: Visualization
12	<ul style="list-style-type: none"> <li>• Write a program which use pandas inbuilt visualization to plot following graphs:               <ol style="list-style-type: none"> <li>a. Bar plots</li> <li>b. Histograms</li> <li>c. Line plots</li> <li>d. Scatter plots</li> </ol> </li> <li>• Write a program to demonstrate use of groupby () method.</li> <li>• Write a program to demonstrate pandas Merging, Joining and Concatenating.</li> <li>• Creating data frames from csv and excel files.</li> </ul>
13.	Lab Exam.
<b>Course Outcomes:</b> At the end of the course student will be able to	
1.	Demonstrate the advanced features of MS Excel
2.	Analyse the data set using the features of MS Excel
3.	Explain the python libraries and mathematical functions for data analysis.

4.	Implement programs in python to extract information from the given data set using suitable libraries.
5.	Apply the visualization functions of MS Excel and Python for visualization of data.

**Course Outcomes Mapping with Program Outcomes & PSO**

Program Outcomes→	1	2	3	4	5	6	7	8	9	10	11	12	PSO↓		
↓ Course Outcomes													1	2	3
<b>AD2601-1.1</b>	3	2	1	-	3	-	-	-	-	-	-	1	1	3	1
<b>AD2601-1.2</b>	3	2	1	-	3	-	-	-	-	-	-	1	1	3	1
<b>AD2601-1.3</b>	3	2	1	-	3	-	-	-	-	-	-	1	1	3	1
<b>AD2601-1.4</b>	3	2	1	-	3	-	-	-	-	-	-	1	1	3	1
<b>AD2601-1.5</b>	3	2	1	-	3	-	-	-	-	-	-	1	1	3	1

**1: Low 2: Medium 3: High**
**REFERENCE BOOKS:**

1.	Wes McKinney, "Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython", O'Reilly, 2nd Edition, 2018
2.	Jake VanderPlas, "Python Data Science Handbook: Essential Tools for Working with Data", O'Reilly, 2017.
3.	Microsoft Excel 2019 Data Analysis and Business Modelling (Business Skills), 6th Edition, Wayne L Winston, ISBN-13: 978-1509305889, ISBN-10: 1509305882.